Introduction to Validation and Uncertainty Quantification

William L. Oberkampf
Consultant
wloconsulting@gmail.com
Austin, Texas

Institute for Computing in Science
Verification, Validation and Uncertainty Quantification Across Disciplines
Park City, Utah
August 6-13, 2011

Outline

- Terminology
- Model accuracy assessment
- Validation metrics
- Steps in non-deterministic simulation
- Concluding remarks

Definition of Error

Error: The difference between an obtained value and the true value.

$$E = y_{obtained} - y_{true}$$

- **The obtained value can be measured, computed, or estimated**
- **Errors can be caused by any anything:**
  - **In experimental measurements: calibration error, random measurement error, systematic measurement error, mistake, blunder, etc**
  - **In modeling: approximations, simplifications, lack of knowledge of physical processes, error in judgment, error in completeness, mistake, blunder, etc**
  - **In simulation: numerical approximations, spatial and temporal discretization, iterative solutions, finite precision arithmetic, programming errors, etc**
- **The true value is only known when certain situations occur:**
  - **Experimental measurements: a calibration standard is used for a quantity**
  - **Modeling and simulation: a highly accurate reference value is used**
- **When a true value is not known or defined, the term uncertainty is more useful than the term error**

Types of Uncertainty

Aleatory uncertainty: uncertainty due to inherent randomness.

- **Also referred to as irreducible uncertainty, variability, and stochastic uncertainty**

Aleatory uncertainty is a characteristic of the system of interest

- **Examples:**
  - **Variation in thermodynamic or mass properties due to manufacturing**
  - **Variation in joint stiffness and damping in structures**

Epistemic uncertainty: uncertainty due to lack of knowledge.

- **Also referred to reducible uncertainty, knowledge uncertainty, model form uncertainty, and subjective uncertainty**

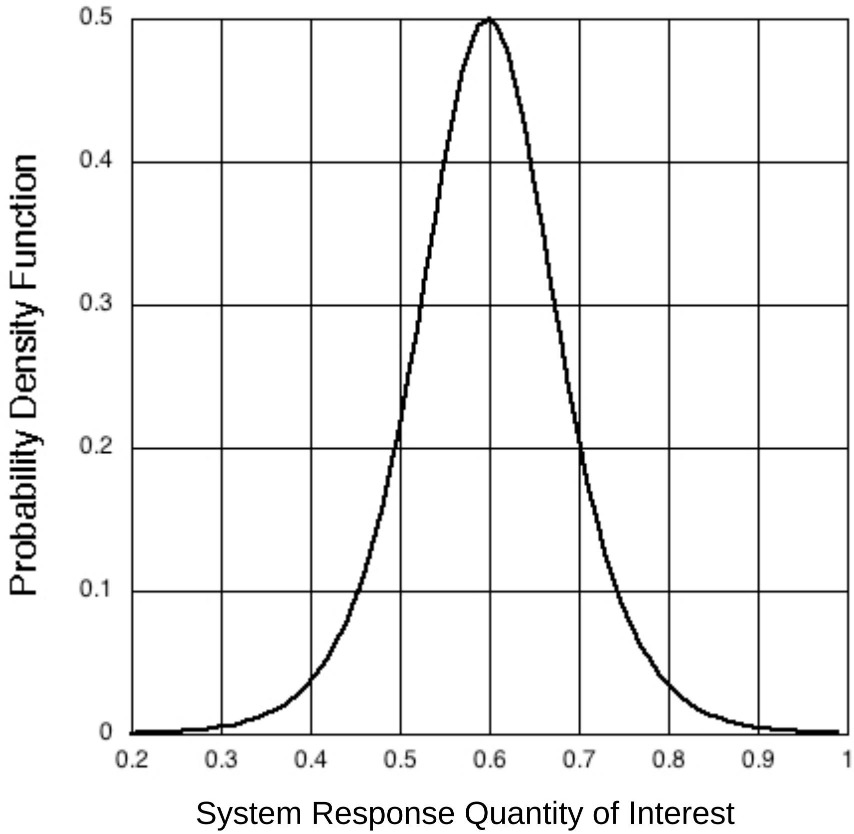
Epistemic uncertainty is a characteristic of our knowledge of the system

- **Examples:**
  - **Poor understanding of physical phenomena, e.g., fracture mechanics**
  - **Poor knowledge or experience of failure, misuse, or hostile scenarios**

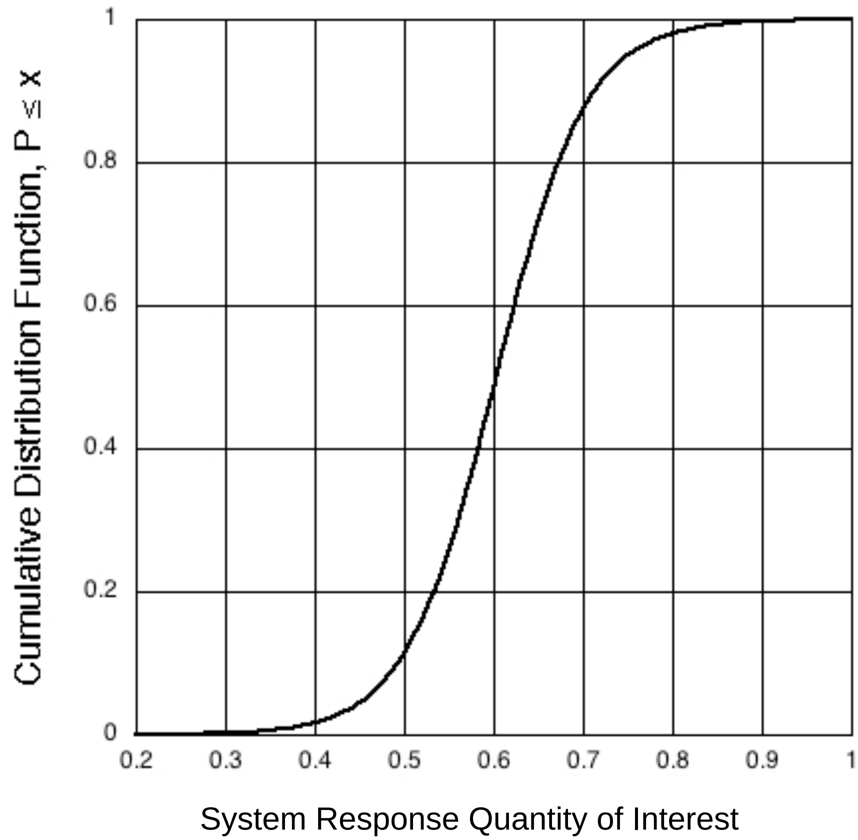
(Ref: Kaplan and Garrick, 1981; Morgan and Henrion, 1990; Ayyub and Klir, 2006)

Characterization of Aleatory Uncertainty

Probability density function
(PDF)

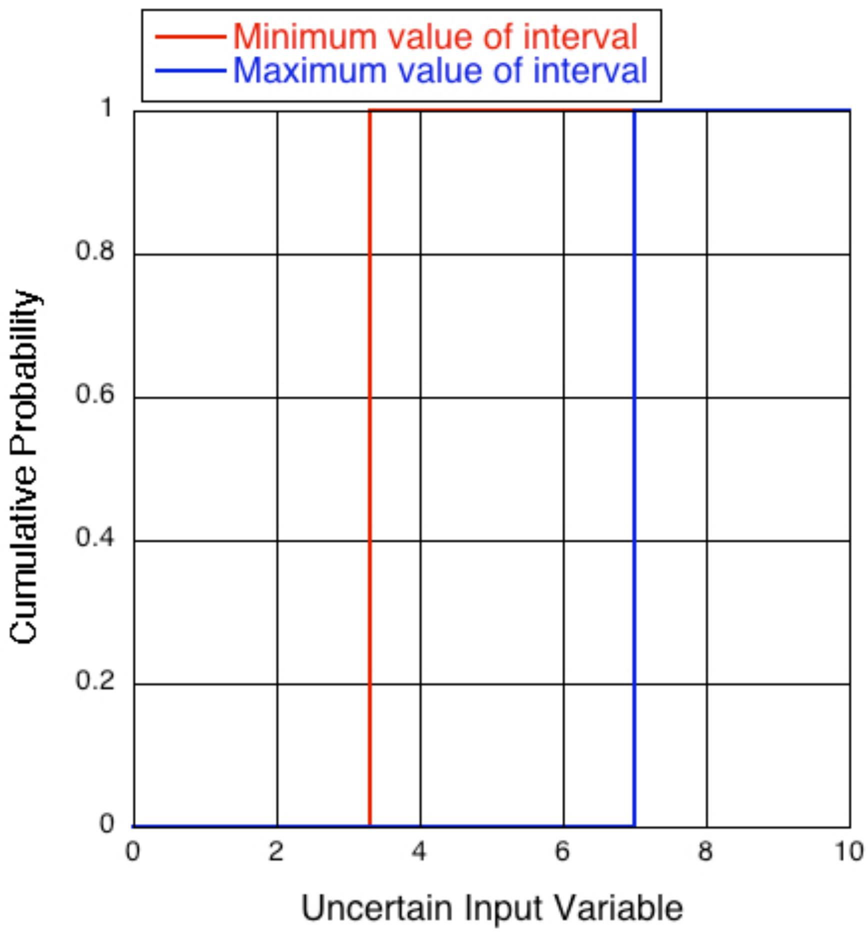


Cumulative distribution function
(CDF)



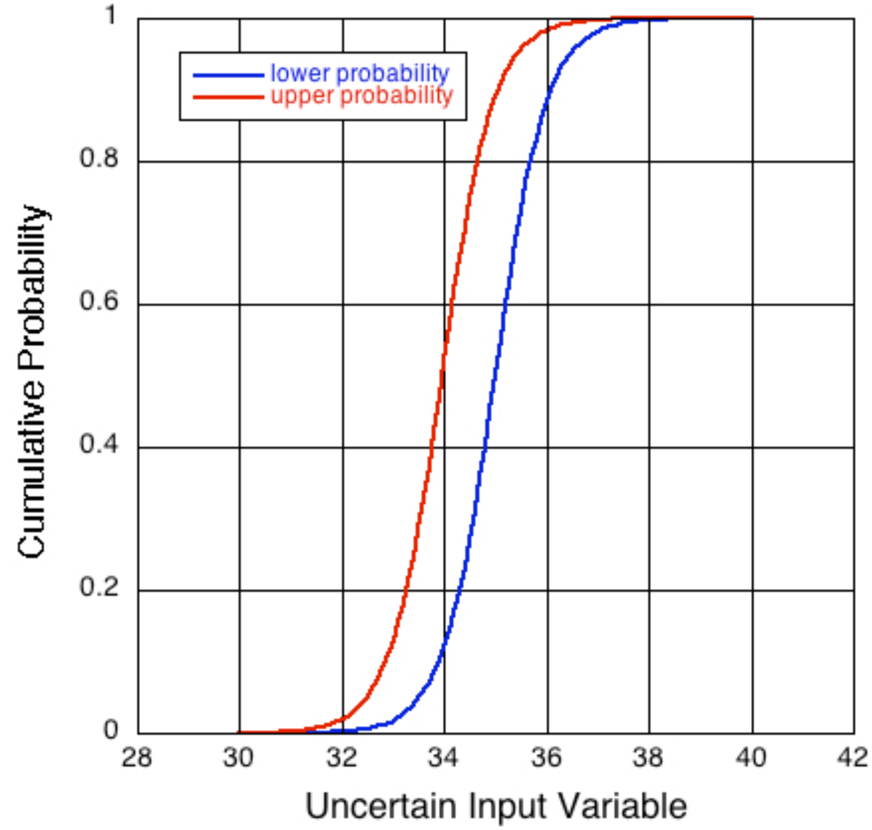
Characterization of Epistemic Uncertainty

A purely epistemic uncertainty is given by an interval (*a*,*b*)



A mixture of epistemic and aleatory uncertainty is given by a p-box

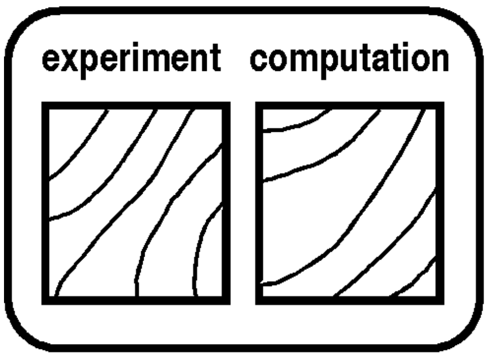
This mathematical structure is referred to as an imprecise probability.



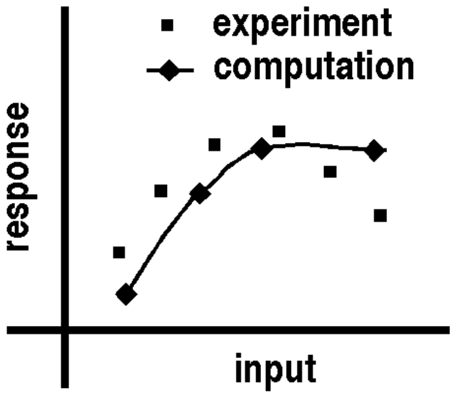
Traditional Methods for Model Accuracy Assessment

Traditional methods of measuring the accuracy of computational results have been either qualitative or semi-quantitative

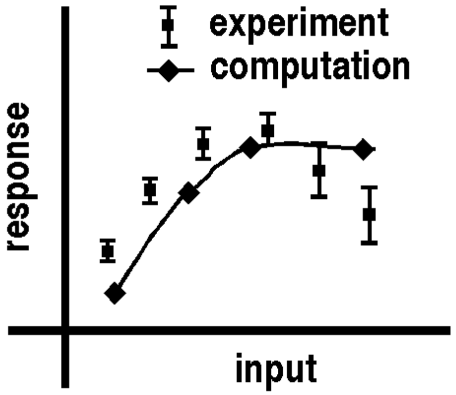
Some examples are:



(a) Viewgraph Norm



(b) Deterministic

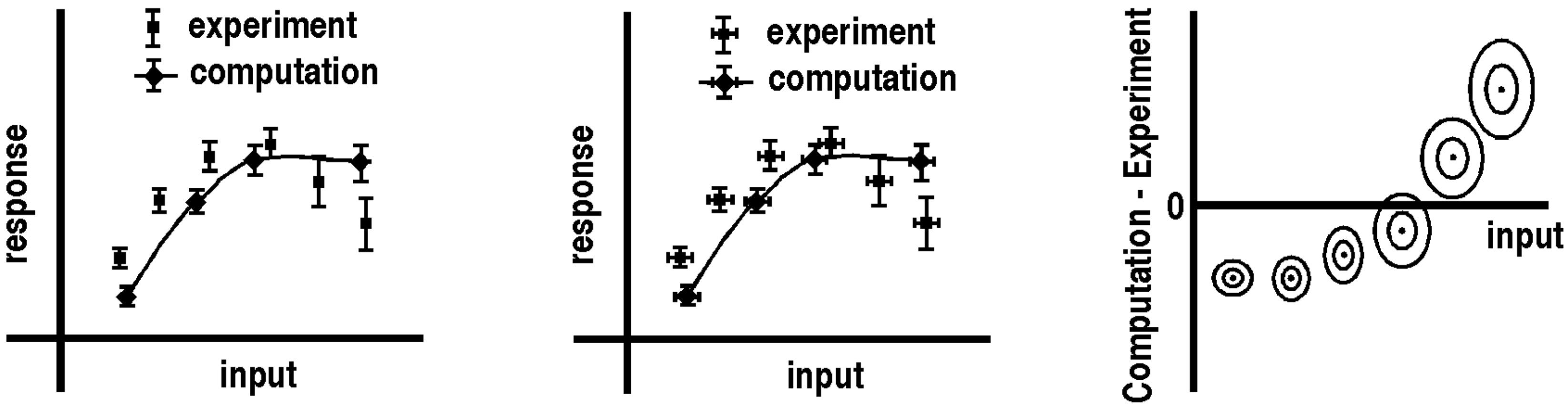


(c) Experimental Uncertainty

(Ref: Oberkampf, Trucano, and Hirsch, 2003)

Traditional Methods for Assessment of Model Accuracy (cont)

More quantitative and precise methods of assessment:



(d) Numerical Error

(e) Nondeterministic Computation

(f) Statistical Comparison

Various Approaches for Comparing Computational and Experimental Results

1) **Estimation of uncertain input parameters to obtain best agreement between simulation and experiment (point estimation):**
   - **Ex: Bendat and Piersol (2000), Wirsching, Paez and Ortiz (1995)**
2) **Hypothesis testing techniques used in statistical inference:**
   - **Ex: Hills and Trucano (2002), Rutherford and Dowding (2003), Dowding et al (2004), Chen et al (2004), Hills (2006)**
   - **The result is a probability that simulation and experiment are "the same"**
3) **Bayesian updating of probability distributions for uncertain input parameters used in the computational model:**
   - **Ex: Kennedy and O'Hagan (2001), Hasselman et al (2001), Zhang and Mahadevan (2003), O'Hagan (2006), Bayarri et al (2007), Chen et al (2008)**
   - **The emphasis is on calibrating probability distributions of parameters**
4) **Comparison of mean values from simulation and experiment:**
   - **Ex: Coleman and Stern (1997), Sprague and Geers (1999), Oberkampf and Trucano (2000), Easterling (2001), Oberkampf and Barone (2006)**
5) **Comparison of cumulative distribution functions from simulation and experiment:**
   - **Ex: Ferson et al (2008), Ferson and Oberkampf (2009)**

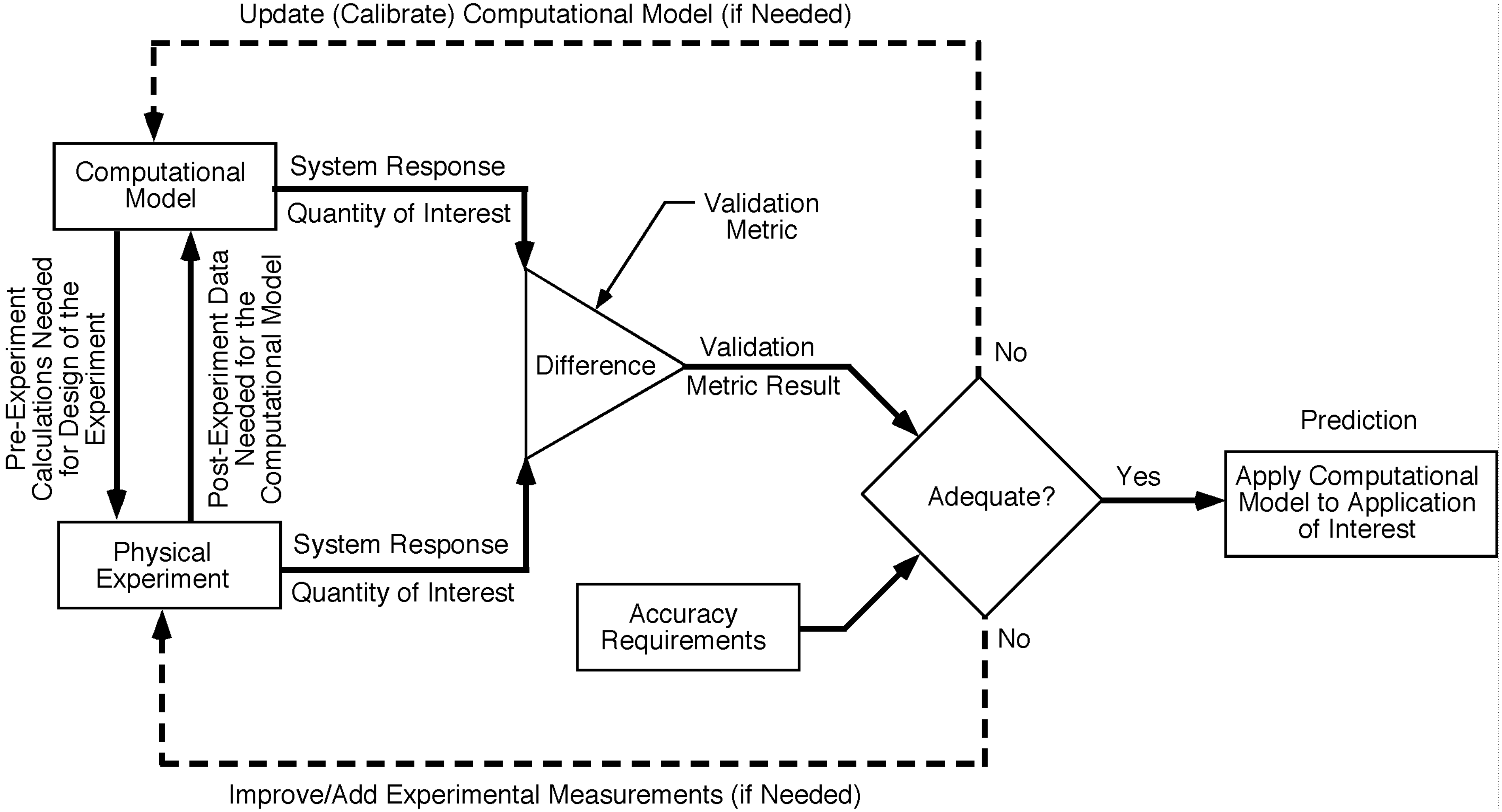
What is a Validation Metric?

- **Validation metric is a statistical measures of agreement between computational results and experimental measurements for system response quantities of interest**

- **Steps to evaluate a validation metric result:**

  1) **Choose a system response quantity of interest**
  2) **Experimentally measure, if possible, all input quantities needed for the code**
  3) **Experimentally measure the system response quantity of interest**
  4) **Using the code and all the input data provided, compute the system response quantity of interest**
  5) **Compute a difference between the experimental measurements and the computational results**

Validation Assessment, Calibration and Prediction

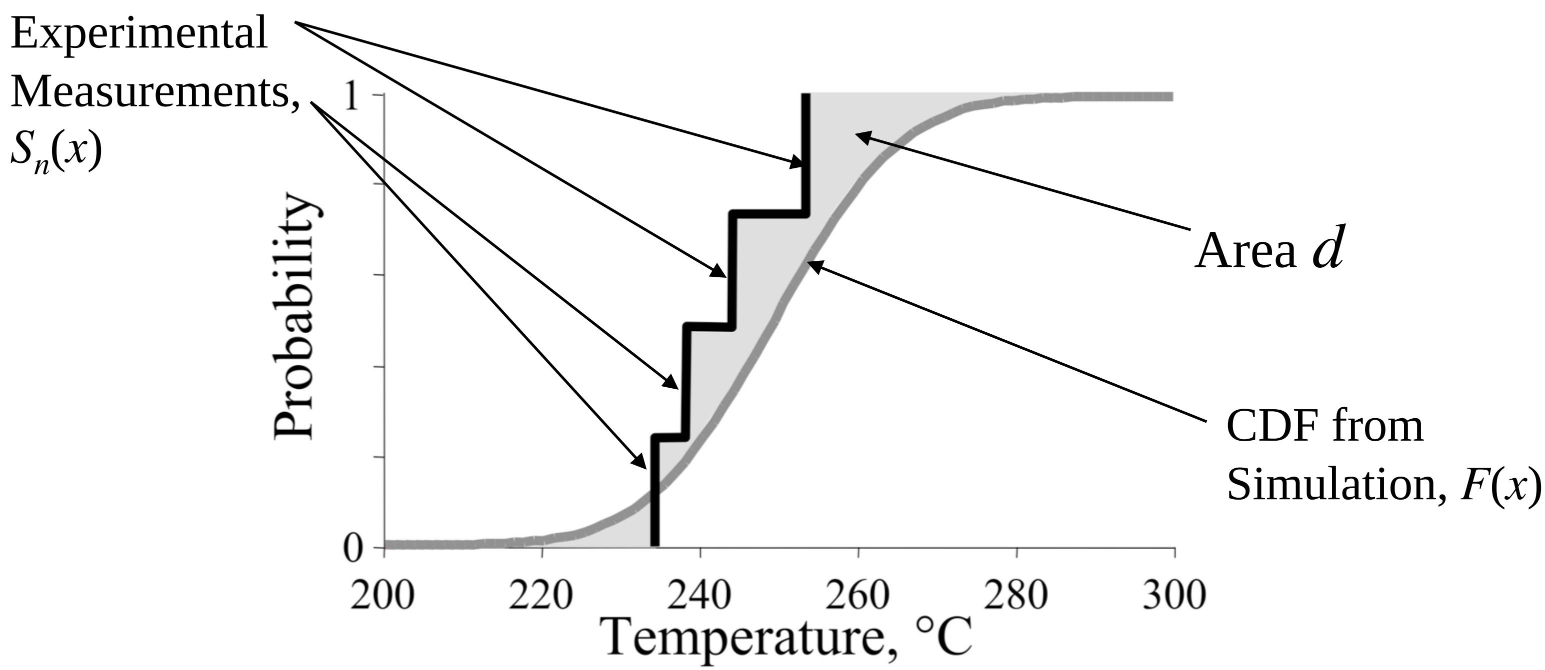


(Ref: Oberkampf and Barone, 2006)

Area Validation Metric

- **The validation metric is defined to be the area between the CDF from the simulation and the empirical distribution function (EDF) from the experiment**

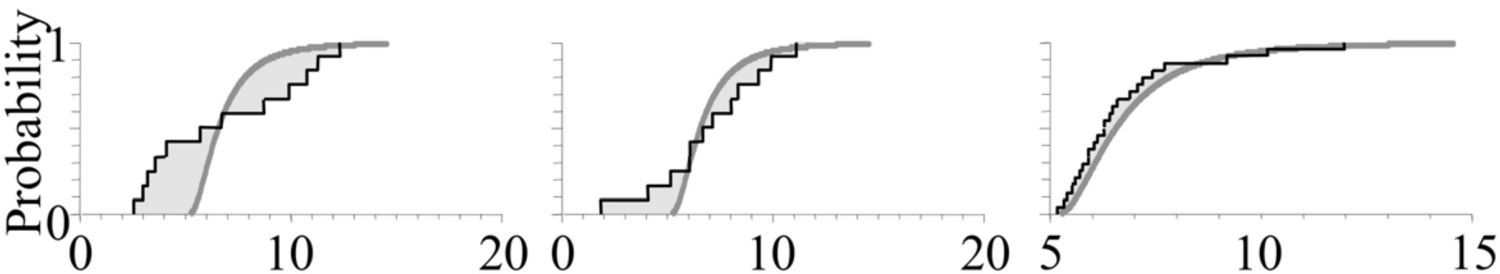
$$d(F, S_n) = \int_{-\infty}^{\infty} \left| F(x) - S_n(x) \right| \mathrm{d}x \qquad \text{(Minkowski } L_1 \text{ metric)}$$



(Ref: Ferson et al, 2008)

Examples of the Area Validation Metric

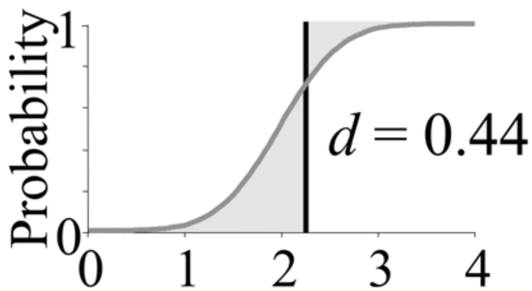
Three different sets of experimental measurements.
The same cumulative distribution function (CDF) predicted by the model.

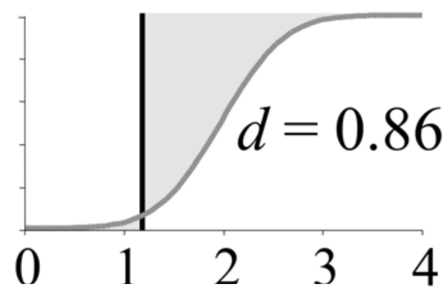


Decreasing values of *d* for each set of measurements.

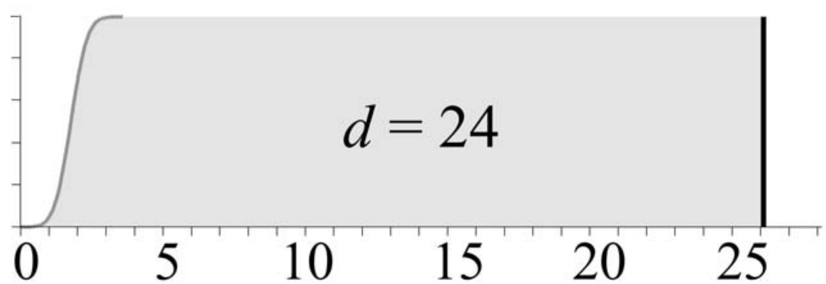
Examples of the Area Validation Metric

A single measurement from three different experiments.
The same cumulative distribution function (CDF) predicted by the model.



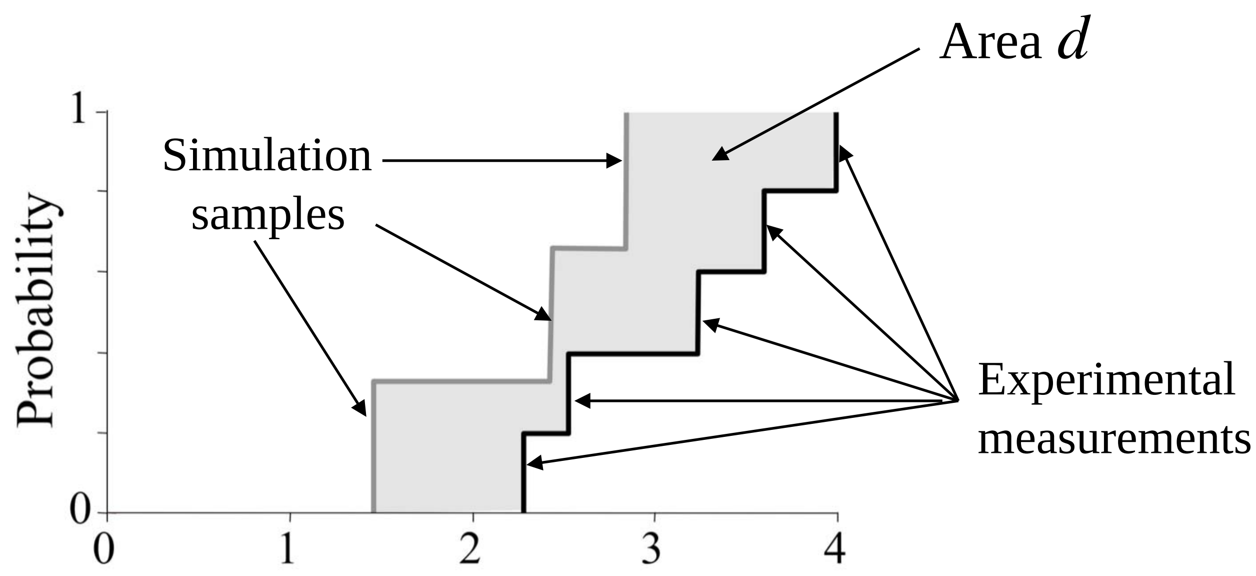


The metric d is dimensional and scale dependent.



Examples of the Area Validation Metric

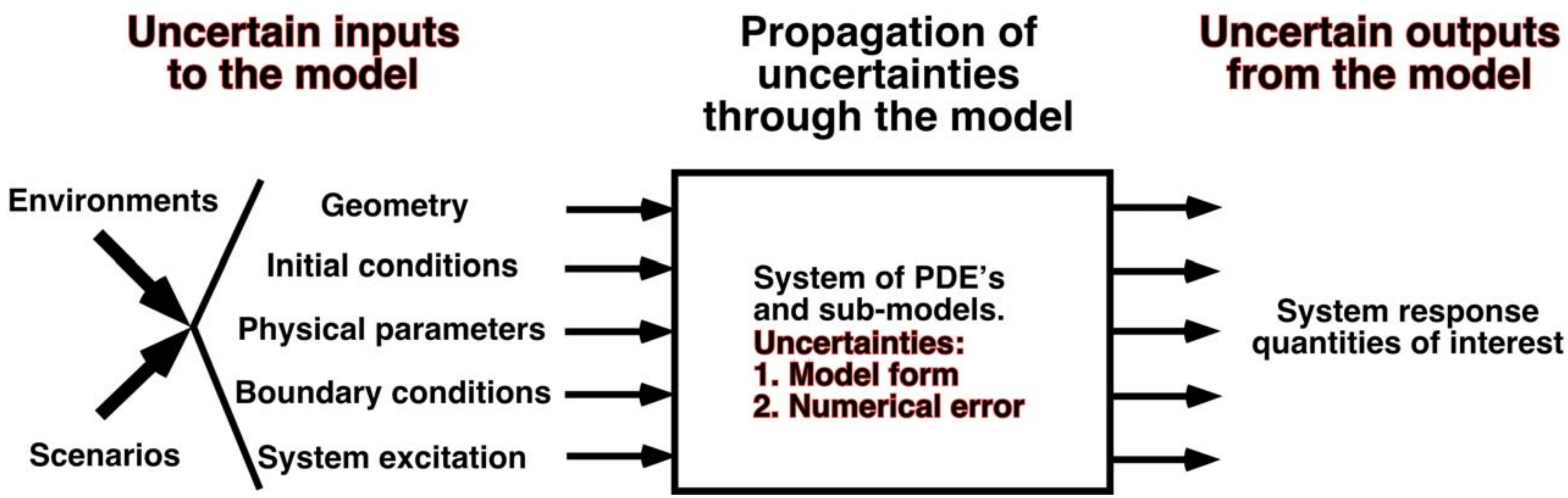
The simulation can also be represented by individual samples.



Features of the area validation metric:

- **The metric measures the shortest average absolute difference of deviates from F and S_n**
- **d = 0 means there is no evidence that the simulation and the experiment are in disagreement.**
- **The area validation metric can also be computed for a probability box, i.e., aleatory and epistemic uncertainty exist in either/both the simulation and the experiment**

Structure of Non-deterministic Simulations

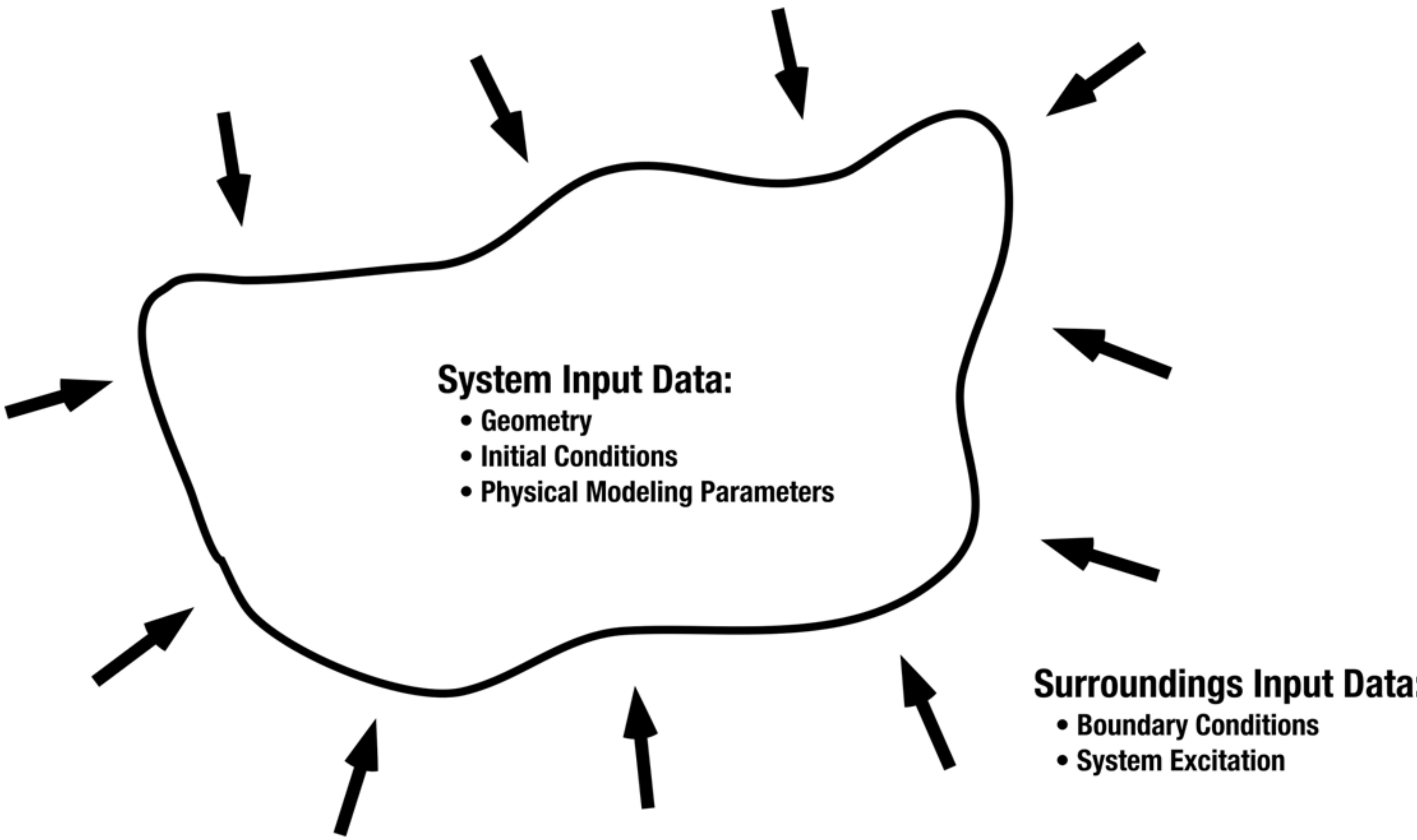


Steps in Non-Deterministic Simulation

1. Identify all relevant sources of uncertainty
2. Characterize each source of uncertainty
3. Estimate solution error in SRQs of interest
4. Estimate uncertainty in SRQs of interest
5. Update model parameters
6. Conduct sensitivity analysis

Step 1: Identify all Relevant Sources of Uncertainty

Identify all model input uncertainties within the system and in the surroundings that can affect system responses.



Sources of Uncertainty

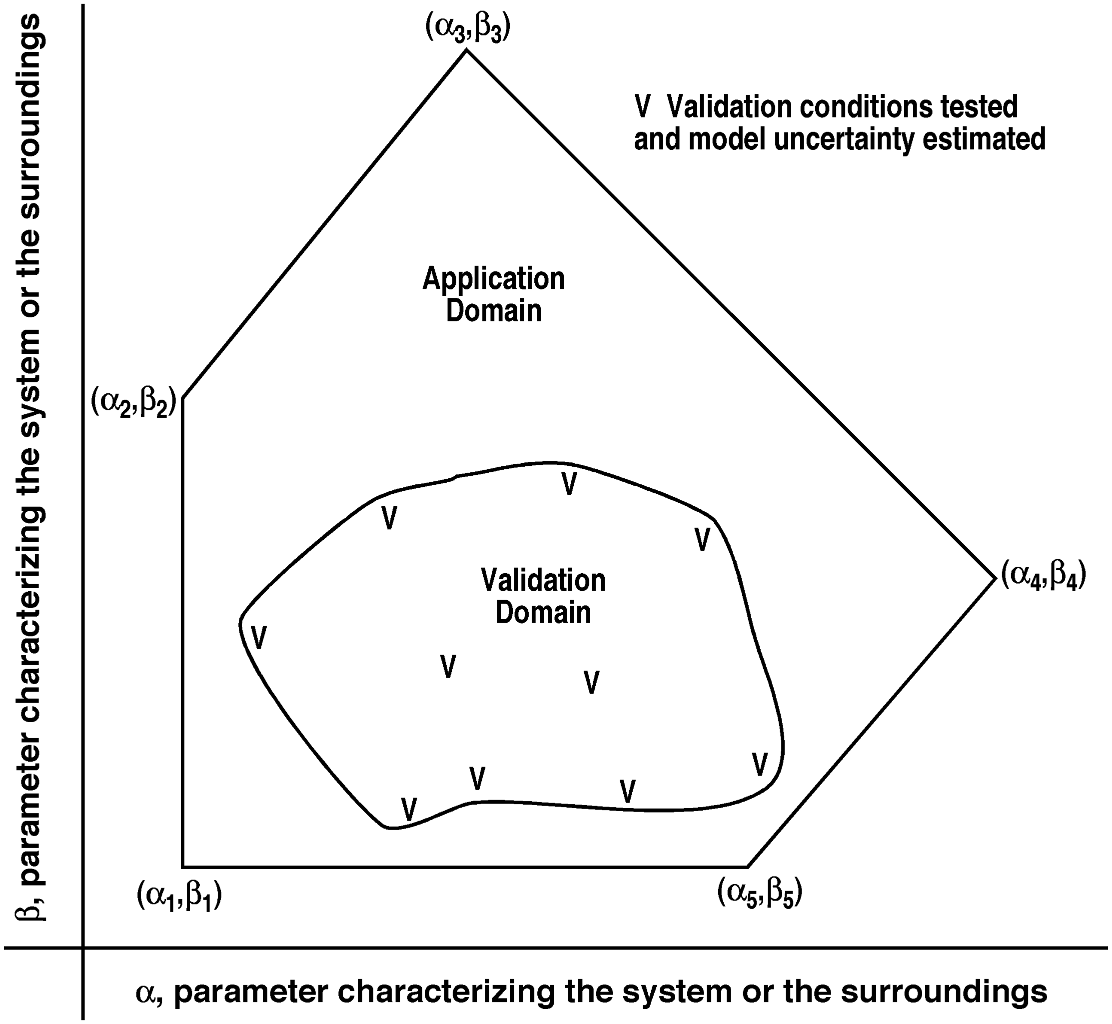
- Uncertainty in model parameters:
  - Input data parameters (independently measureable and non-measureable)
  - Uncertainty modeling parameters
  - Numerical algorithm parameters
- Numerical solution error:
  - Round-off error
  - Iterative error
  - Spatial and temporal discretization error
- Model form uncertainty:
  - Estimated over the validation domain
  - Extrapolated outside of the validation domain

Step 2: Characterize Each Source of Uncertainty

There are three dominant approaches to nondeterministic simulation:

1. **Traditional probabilistic methods**
   - All uncertainties are characterized as a probability distribution
   - Epistemic (lack of knowledge) uncertainties are either ignored or characterized as a uniform probability density function
2. **Bayesian inference**
   - Assume prior distributions represent both frequency of occurrence (aleatory) uncertainties and personal belief of likelihood (epistemic) uncertainties
   - Update the prior distributions for uncertain parameters using available experimental data and Bayes formula
   - Compute new predictions using updated parameter distributions
3. **Probability bounds analysis**
   - Closely related to two-dimensional (or second order) Monte Carlo methods
   - Keep aleatory and epistemic uncertainties segregated throughout the analysis
   - Characterize aleatory uncertainties as probability distributions
   - Characterize epistemic uncertainties as interval-valued quantities
   - Represent system response quantities (SRQs) as interval-valued probability distributions, i.e., p-boxes

Model Uncertainty



- **To estimate model uncertainty, a validation metric result must be computed over the application domain**
- **Over the validation domain, one can use either an:**
  - **Interpolation function**
  - **Regression fit**
- **Beyond the validation domain, one must extrapolate the validation metric**

Step 3:
Estimate Solution Error in SRQs of Interest

Estimate the magnitude of the numerical solution errors on each of the SRQs of interest.

(Code verification should have been completed on all code options that are exercised in the analysis.)

- **Use stationary and non-stationary iterative error estimators for:**
  - **Initial value problems (stationary methods commonly used)**
  - **Boundary value problems (stationary and Krylov subspace methods used)**
- **Two methods for controlling temporal discretization error:**
  - **Error estimated at each time step, compared to some error criterion, and appropriate adjustments are made in the time step size**
  - **An entire solution is computed with a fixed time step and then recomputed with either a smaller or larger time step**

Estimation of Spatial Discretization Error

- **Zienkiewicz-Zhu super-convergent patch recovery (ZZ-SPR) method is probably the most widely used finite-element-based error estimator.**
- **ZZ-SPR can provide error estimates in local SRQs if:**
  - **Finite element types are used that have the super-convergent property**
  - **There is strong evidence that the mesh is adequately resolved**
- **Richardson-extrapolation-based methods can be used on essentially any SRQs if:**
  - **Meshes are uniformly refined from one mesh to the next**
  - **Meshes are in the asymptotic region**
  - **A minimum of two mesh solutions are required, but this is risky**

The sum of all numerical solution error contributors can be written as

$$(U_N)_{y_i} = \left|U_I\right|_{y_i} + \left|U_S\right|_{y_i} + \left|U_T\right|_{y_i} \quad for \quad i = 1, 2, \ldots n$$

where each quantity is an interval-valued quantity, i.e., an epistemic uncertainty

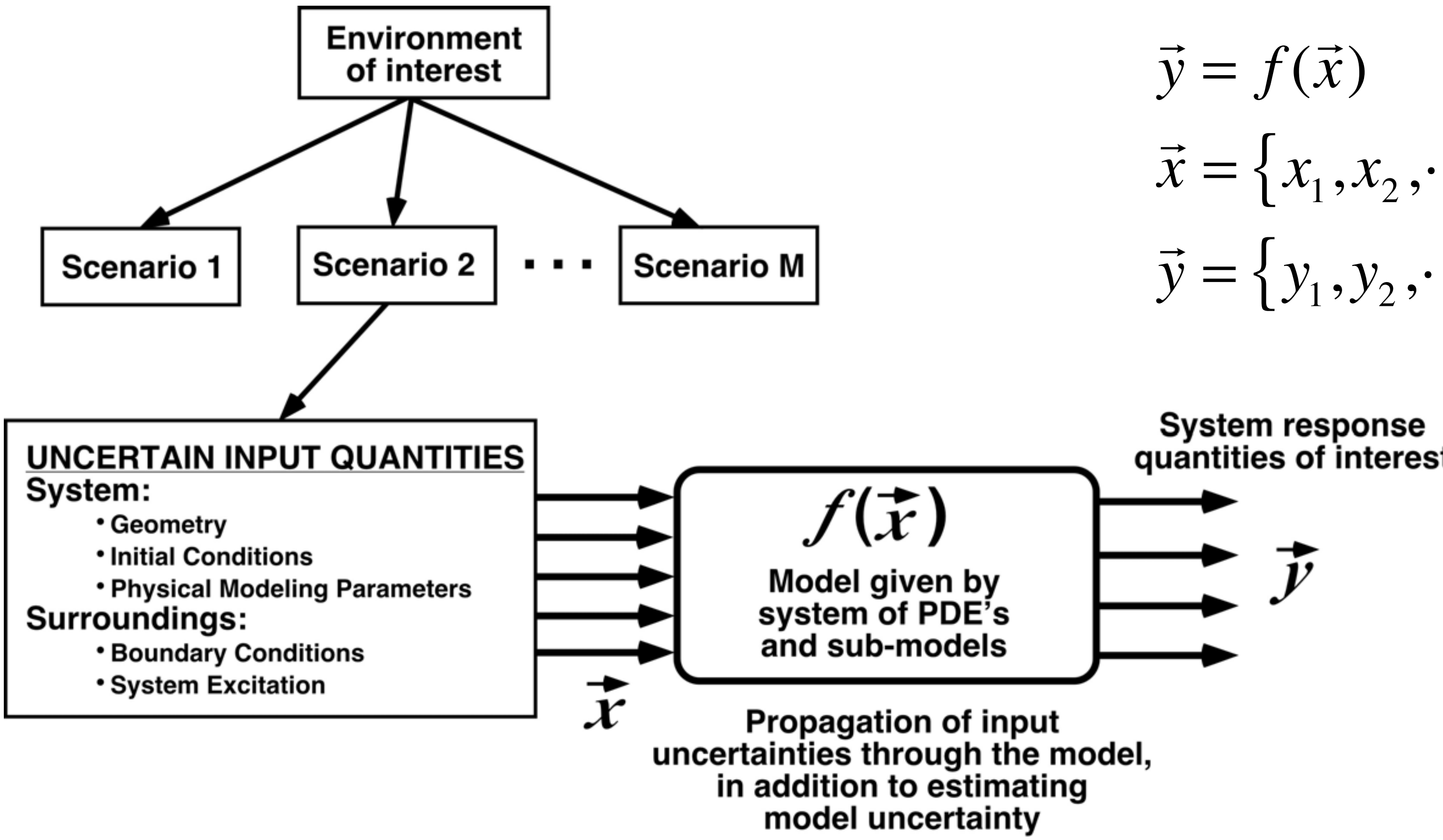
Step 4:
Estimate Uncertainty in SRQs of Interest

Propagate all sources of uncertainty through the through the model to obtain the uncertainty in the SRQs of interest.

$$\vec{y} = f(\vec{x})$$

$$\vec{x} = \{x_1, x_2, \cdots x_m\}$$

$$\vec{y} = \{y_1, y_2, \cdots y_n\}$$



Example of the Propagation Technique for Probability Bounds Analysis

- **Let $\vec{x}$ be decomposed into aleatory and epistemic input quantities**

$$\vec{x} = \left(\vec{x}_A, \vec{x}_E\right)$$

  - **where $\vec{x}_A = x_1, x_2, \cdots x_\alpha$ represents all aleatory uncertainties**
  - **where $\vec{x}_E = x_{\alpha+1}, x_{\alpha+2}, \cdots x_m$ represents all epistemic uncertainties**

- **Epistemic uncertainties typically occur in environments, scenarios, physical parameters, and parameters in ICs, BCs, and system excitation**

- **Using PBA and sampling methods (Monte Carlo or Latin Hypercube), one can keep aleatory and epistemic uncertainties separated**

- **Use a sampling procedure with an inner loop for the aleatory uncertainties, and an outer loop for the epistemic uncertainties:**

outer

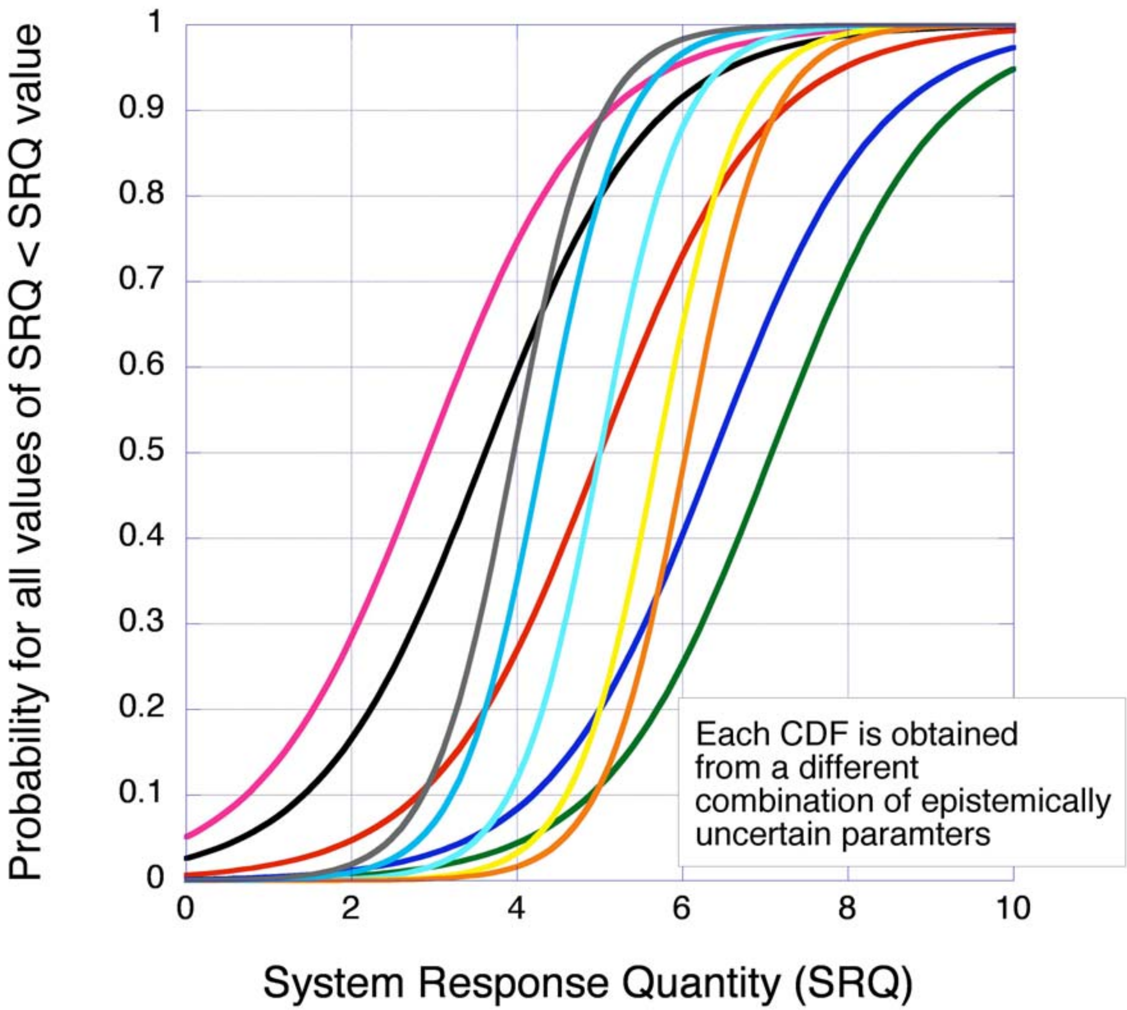
inner

Choose one sample from all x_i in $\vec{x}_E$

- **Choose a random sample from each x_i in $\vec{x}_A$**
  - **Propagate sampled quantities through the model**
- **Choose random samples from $\vec{x}_A$ until a satisfactory CDF is obtained**

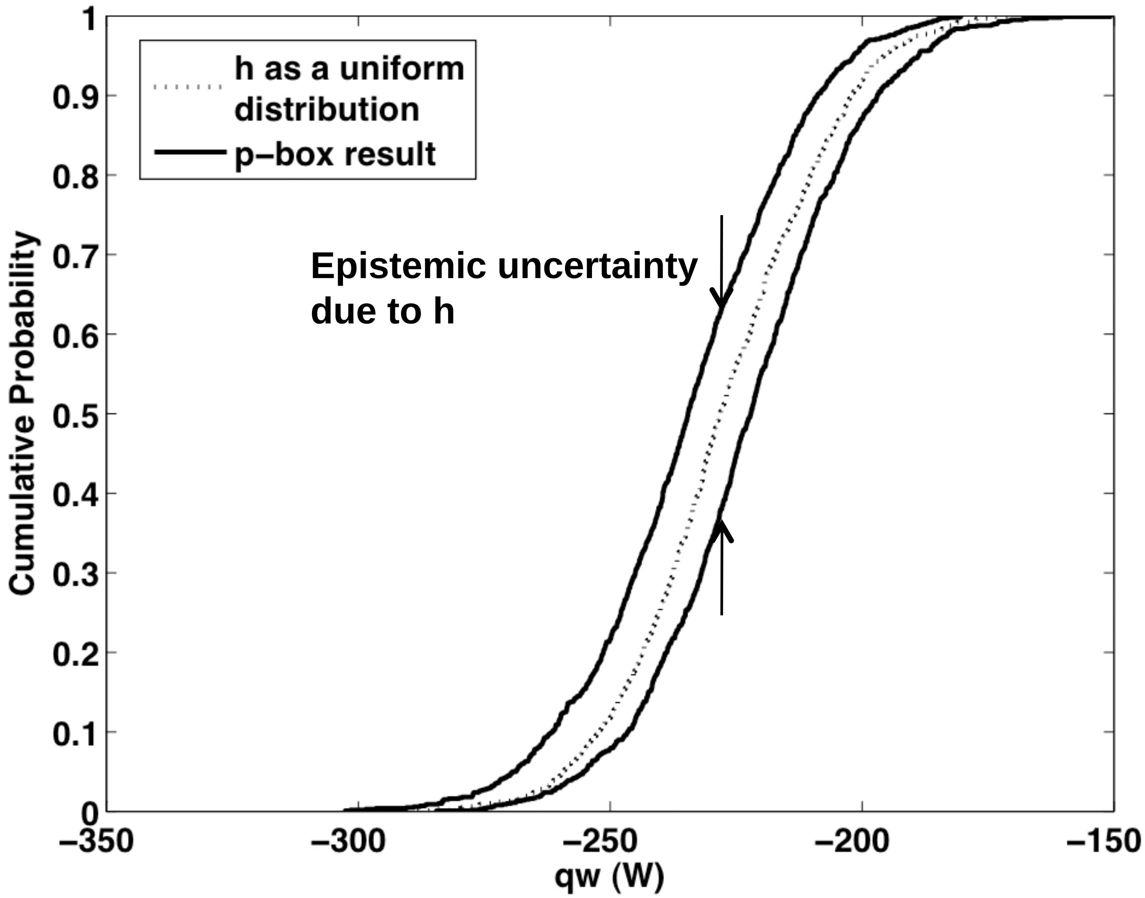
Choose random samples from all $\vec{x}_E$ until sufficient converge is obtained

Cumulative Distribution Functions for an SRQ of Interest



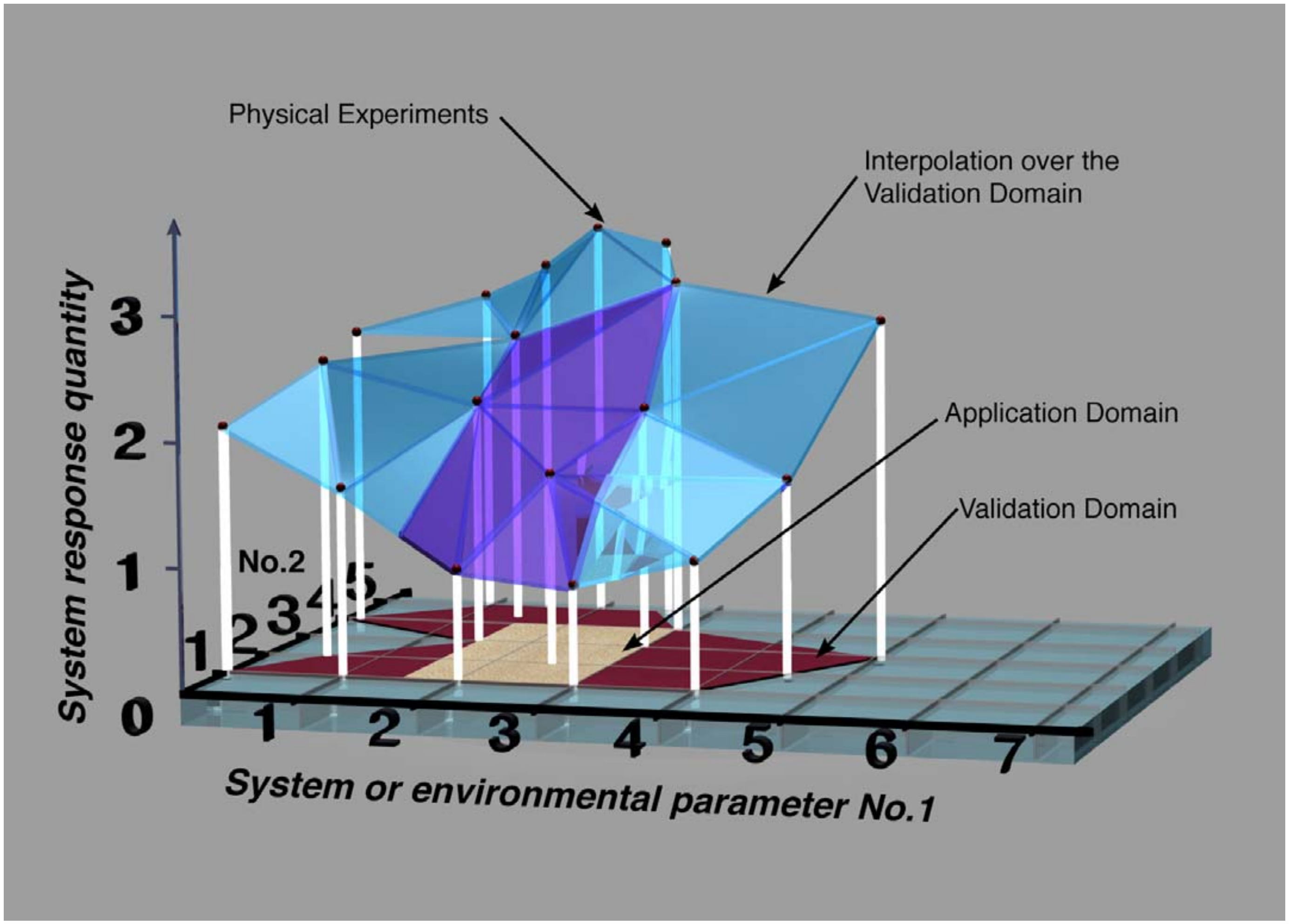
- **For each sample of $\vec{x}_E$, a number of samples of $\vec{x}_A$ are propagated through the model**
- **Here we have ten samples for $\vec{x}_E$**
- **All probabilities between the minimum and maximum CDFs are possible at a given value of SRQ**
- **The result is an interval valued probability, a p-box**

Example of a p-box Due to Aleatory and Epistemic Input Uncertainty



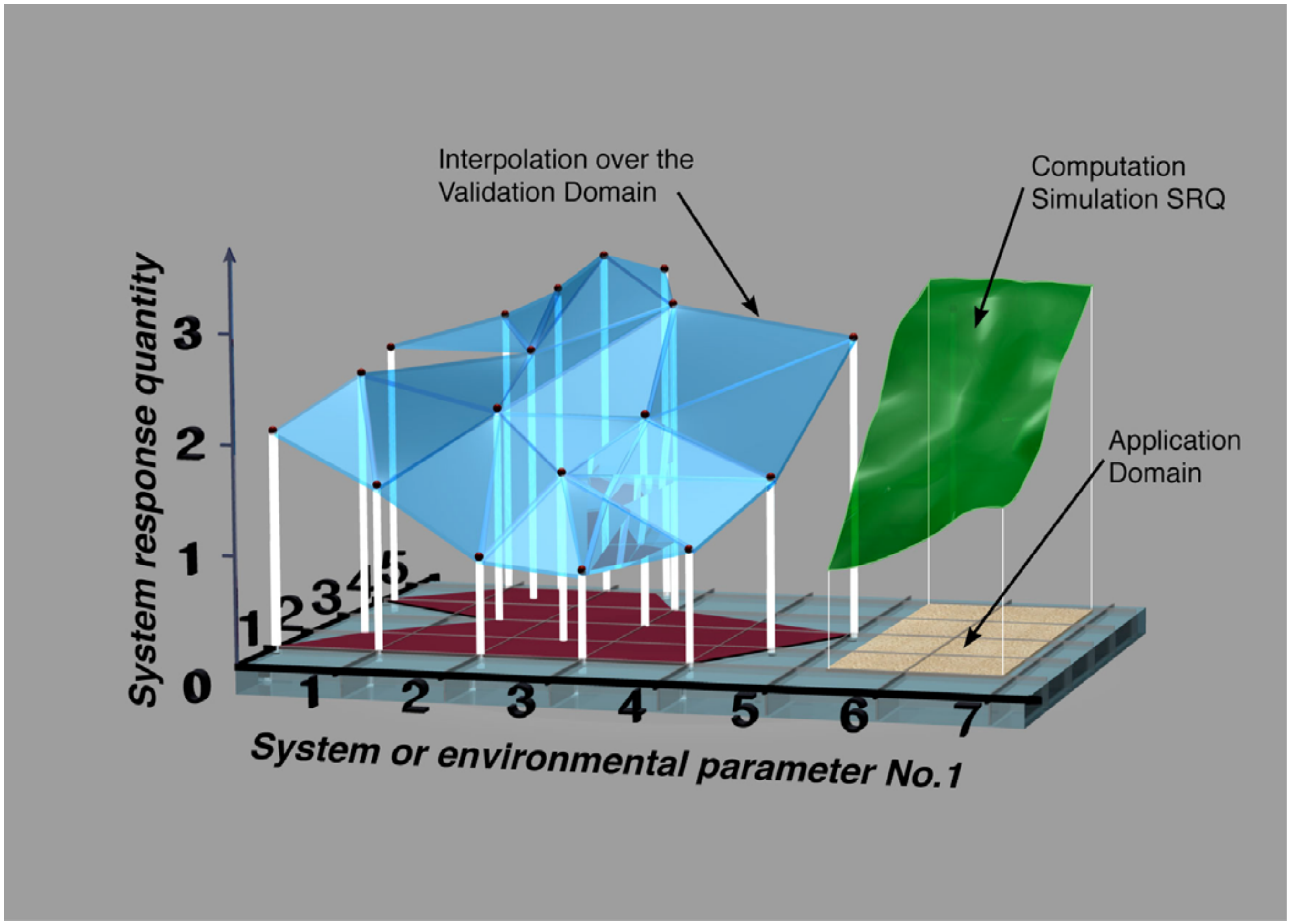
- Example of 2-D, steady state, heat transfer through a wall (Ch. 13)
- Thermal conductivity, k, is an aleatory uncertainty
- Convective heat transfer coefficient, h, is an epistemic uncertainty
- If h is treated as a uniform PDF in a traditional nondeterministic analysis, the uncertainty is under represented

Prediction Within the Validation Domain: Interpolation



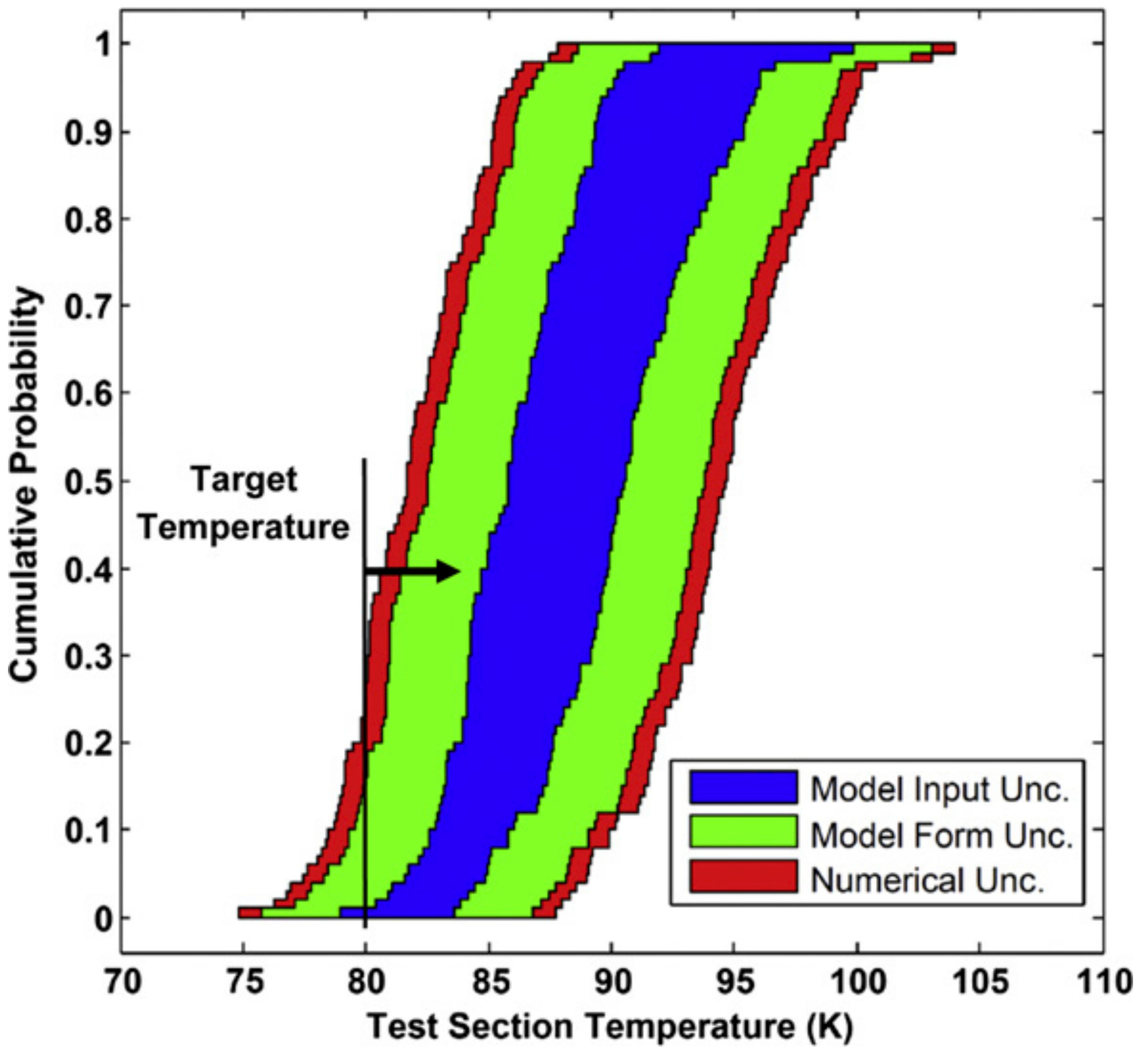
- **Prediction can be:**
  - **Interpolated between validation points**
  - **Computed using surrogate models**
- **A validation metric can be computed at each experimental point**
- **Difficult to determine if you are interpolating or extrapolating in a high-dimensional input space**

Prediction Far Outside the Validation Domain: Large Extrapolation



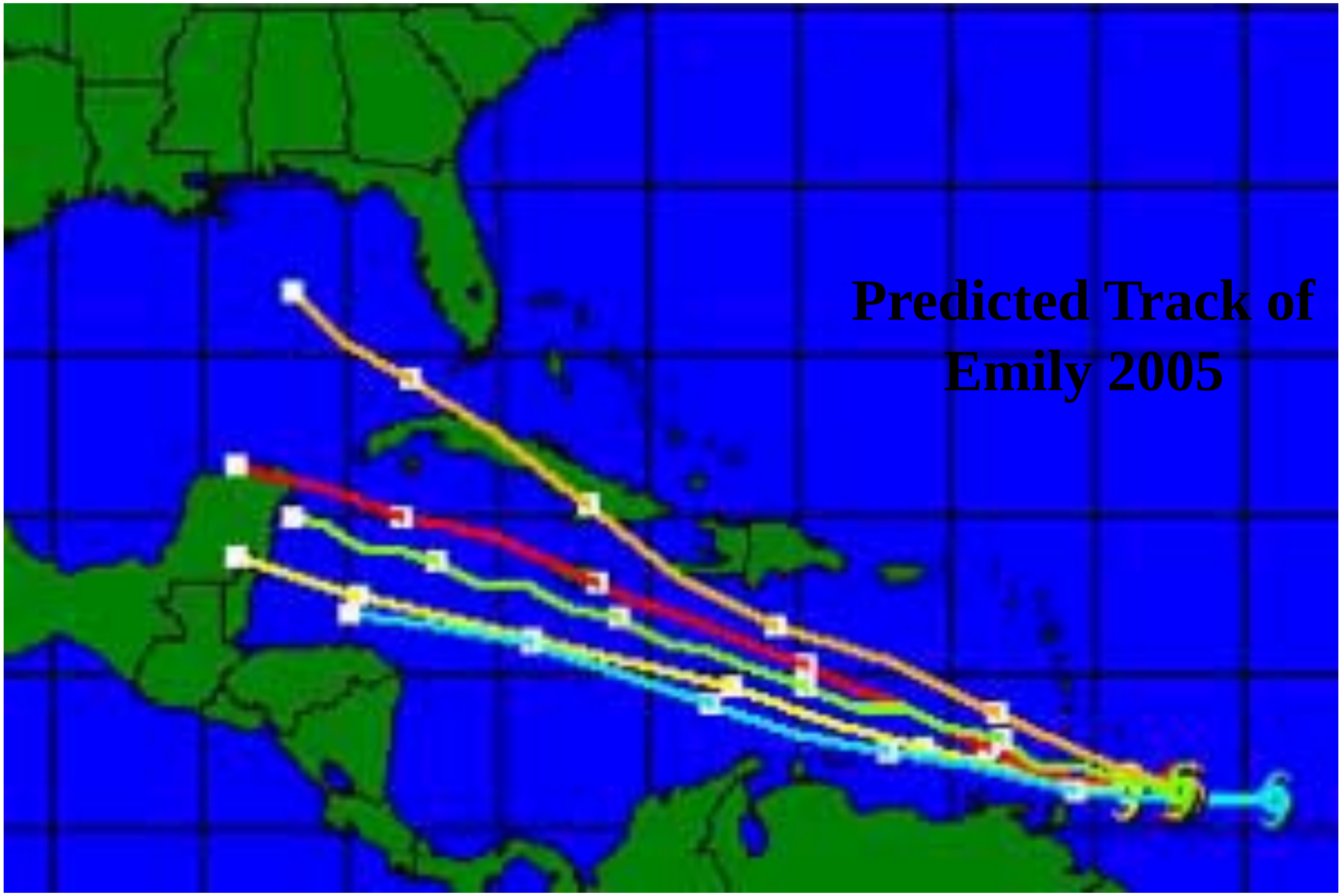
- **Extrapolations can occur in terms of:**
  - **Input quantities**
  - **Model uncertainty**
- **Extrapolation may also require:**
  - **Large changes in coupled physics, e.g., heating effects on structural dynamics**
  - **Large changes in geometry or subsystem interactions, e.g., abnormal or hostile environments**
- **Large extrapolations should result in large increases in uncertainty**

Example of p-box with a Mixture of Aleatory and Epistemic Uncertainty



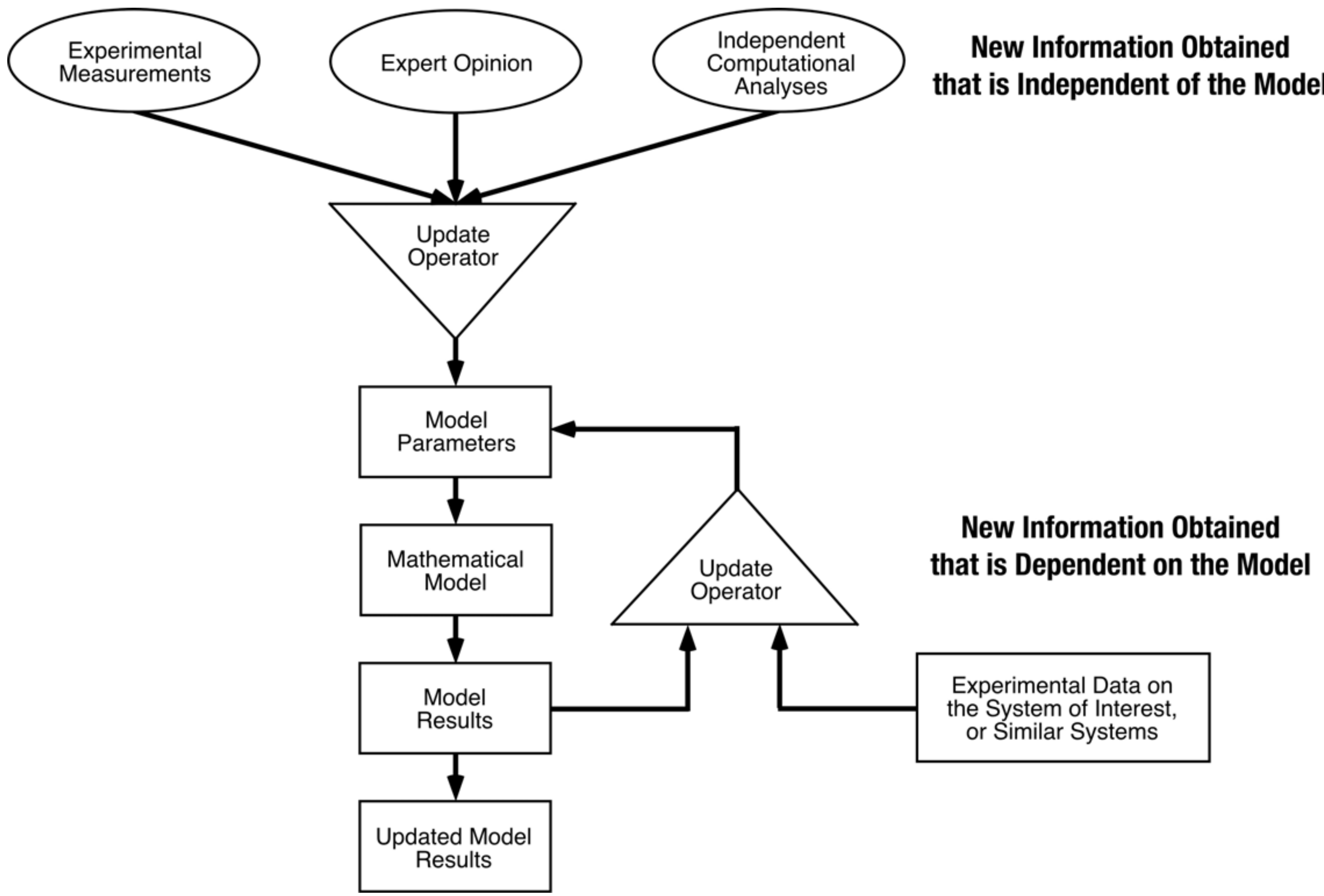
From Roy and Oberkampf (2011)

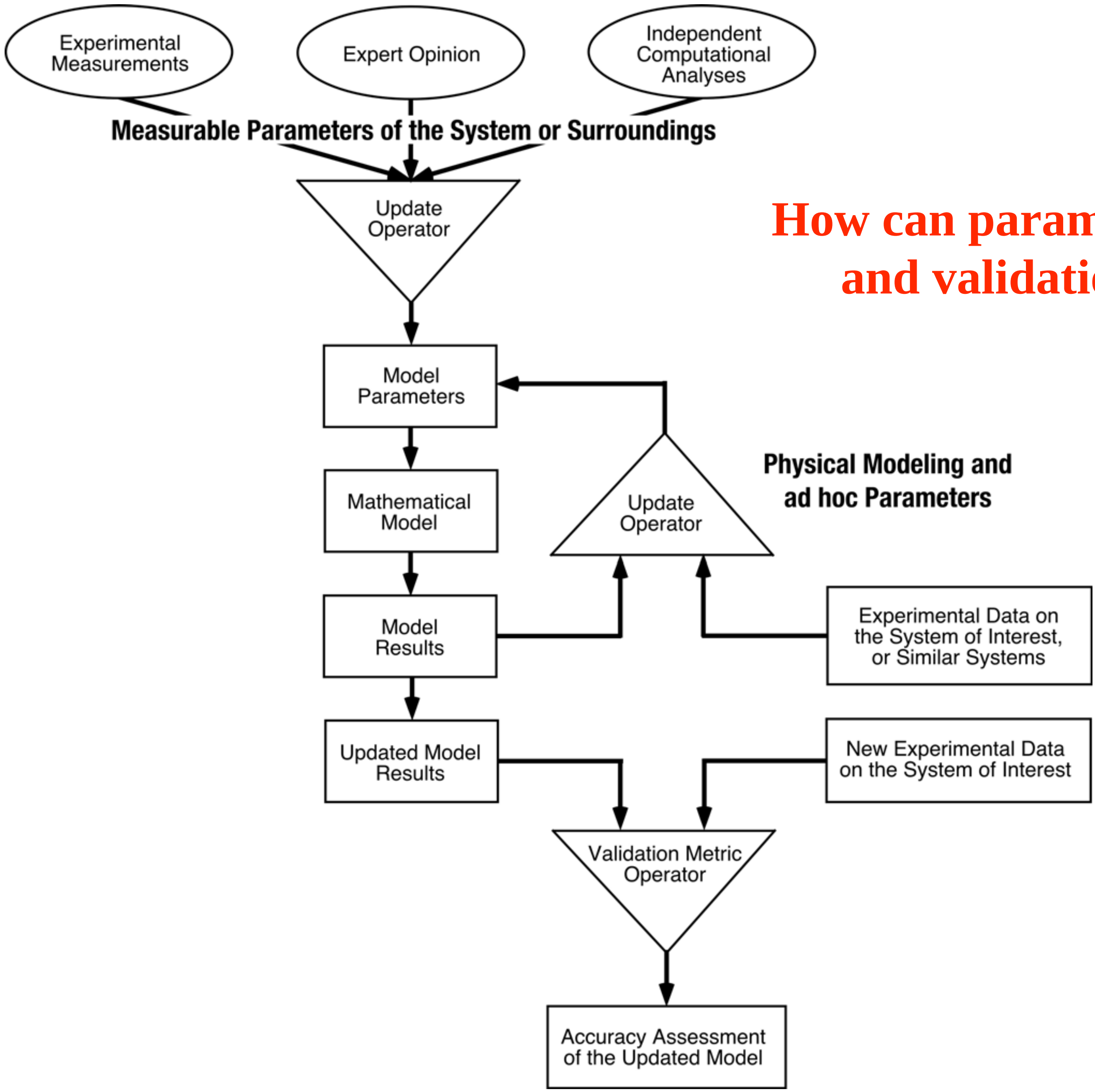
Example Showing Model Form Uncertainty



From Green (2007)

Step 5:
Update Model Parameters





How can parameter updating and validation coexist?

Step 6:
Conduct a Sensitivity Analysis

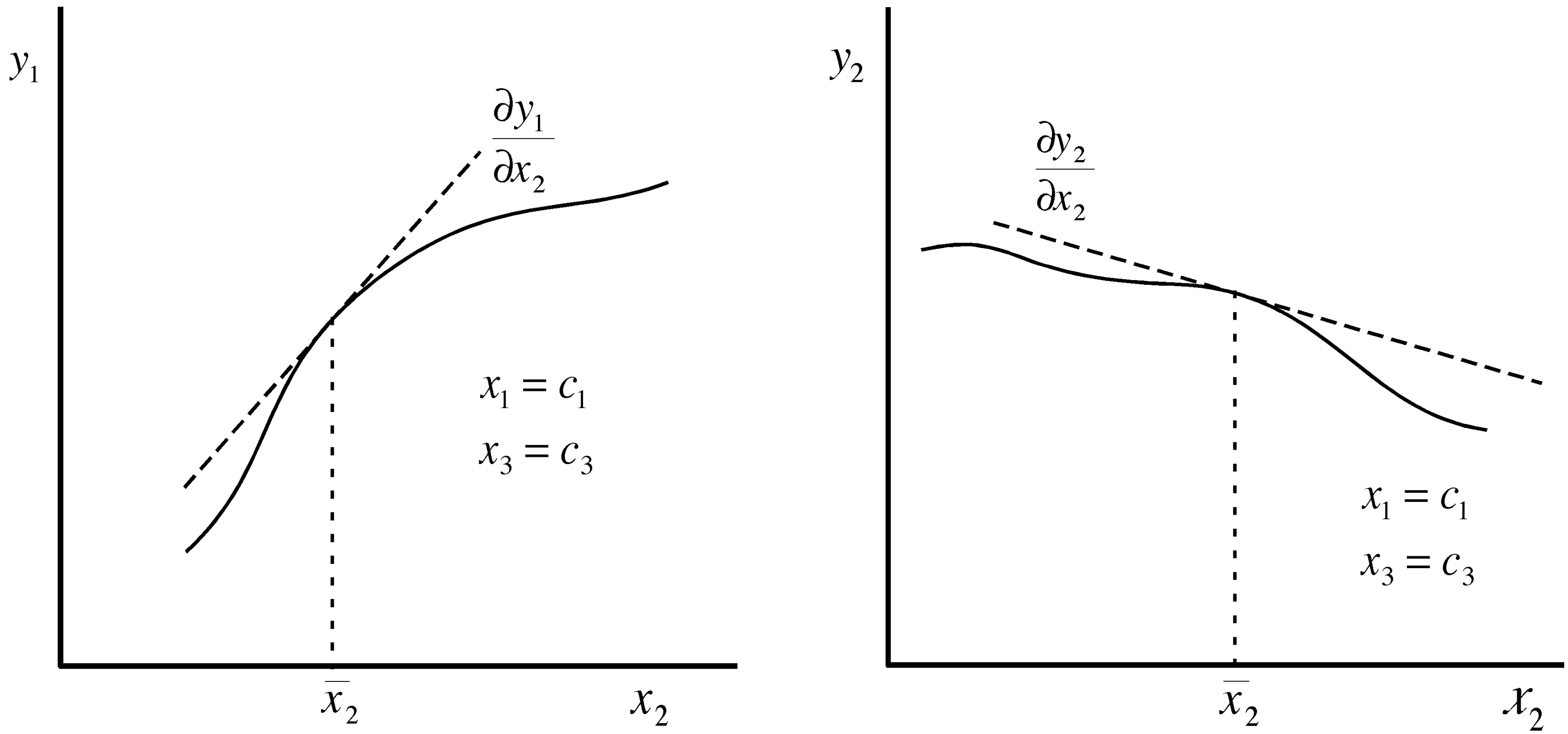
Sensitivity analysis (SA) is the determination how a change in any aspect of the model changes any response of the model.

- **Local SA determines how outputs change locally as a function of inputs:**
  - **Commonly used to determine which system design parameters can be optimized for improved system performance or safety**
  - **Input parameters are optimized for specific conditions, e.g., expected value**
- **Global SA determines how the uncertainty structure of the inputs maps to the uncertainty structure of the outputs:**
  - **Analysis usually begins with examining scatter plots of responses as a function of input uncertainties**

(Helton et al, 2006; Saltelli et al, 2008)

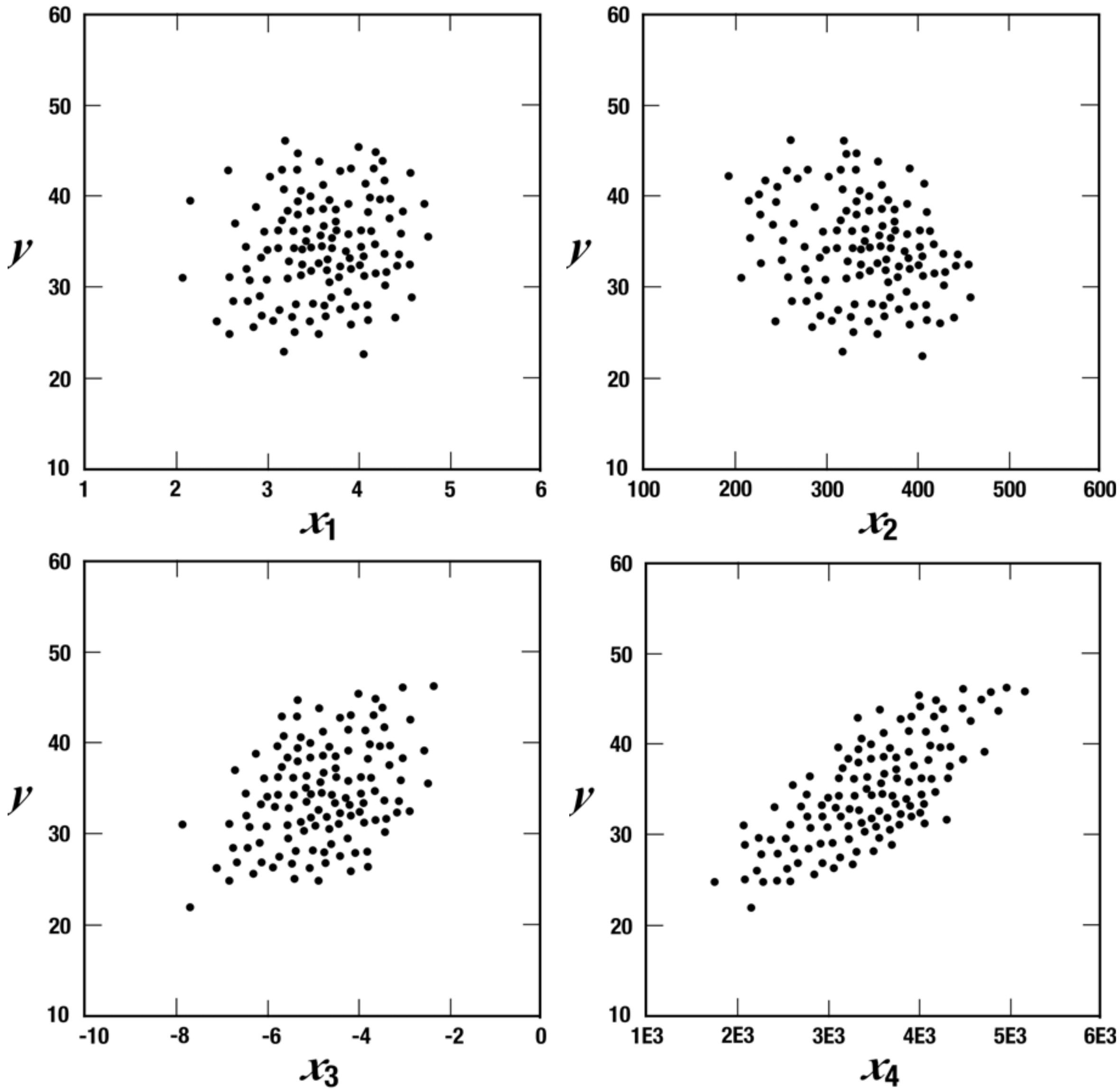
Example of Local Sensitivity Analysis

Suppose that $\vec{x} = \{x_1, x_2, x_3\}$ and $\vec{y} = \{y_1, y_2\}$



Numerical solution "noise" can arise and yield erroneous results

Example of Global Sensitivity Analysis



- **Assume that** $y = f\left(x_1, x_2, x_3, x_4\right)$
- **All uncertain inputs are aleatory**
- **Assume that 100 Monte Carlo samples have been computed**
- **Then one can determine if any trend in** y **exists as a function of** x_1, x_2, x_3, x_4

Concluding Remarks

- Validation metrics are proving very effective in quantifying model accuracy
- Bayesian updating:
  - Traditional approaches are model calibration
  - Kennedy and O'Hagan's (2001) approach calibrates parameters and attempts to quantify model accuracy at the same time
- Ignoring numerical solution error and model form uncertainty will underestimate the total predictive uncertainty
- Additional research is needed to improve extrapolation of model form uncertainty
- Simulation of chaotic processes is highly suspect
- How much validation and uncertainty quantification is enough?

References

- **Ayyub, B. M. and G. J. Klir (2006). Uncertainty Modeling and Analysis in Engineering and the Sciences, Boca Raton, FL, Chapman & Hall.**
- **Bayarri, M. J., J. O. Berger, R. Paulo, J. Sacks, J. A. Cafeo, J. Cavendish, C. H. Lin, and J. Tu (2007), "A Framework for Validation of Computer Models,"** ***Technometrics*****, Vol. 49, No. 2, pp. 138-154.**
- **Bendat, J. S. and A. G. Piersol, (2000), Random Data: Analysis & Measurement Procedures, 3rd Edition, Wiley, New York.**
- **Coleman, H. W. and F. Stern (1997), "Uncertainties and CFD Code Validation,"** ***Journal of Fluids Engineering*****, Vol. 119, pp. 795-803.**
- **Chen, W., L. Baghdasaryan, T. Buranathiti, and J. Cao (2004), "Model Validation via Uncertainty Propagation and Data Transformations,"** ***AIAA Journal*****, Vol. 42, No. 7, pp. 1406-1415.**
- **Chen, W., Y. Xiong, K-L Tsui, and S. Wang (2008), "A Design-Driven Validation Approach Using Bayesian Prediction Models,"** ***Journal of Mechanical Design*****, Vol. 130, No. 2.**
- **Dowding, K., R. G. Hills, I. Leslie, M. Pilch, B. M. Rutherford, and M. L. Hobbs (2004), "Case Study for Model Validation: Assessing a Model for Thermal Decomposition of Polyurethane Foam," Sandia National Laboratories, SAND2004-3632, Albuquerque, NM.**
- **Easterling, R. G. (2001), "Measuring the Predictive Capability of Computational Models: Principles and Methods, Issues and Illustrations," Sandia National Laboratories, SAND2001-0243, Albuquerque, NM.**

References (continued)

- **Ferson, S., W. L. Oberkampf, and L. Ginzburg (2008), "Model Validation and Predictive Capability for the Thermal Challenge Problem," *Computer Methods in Applied Mechanics and Engineering*, Vol. 197, pp. 2408-2430.**
- **Ferson, S. and W. L. Oberkampf (2009), "Validation of Imprecise Probability Models," *International Journal of Reliability and Safety*, Vol. 3, No. 1-3, pp. 3-22.**
- **Ferson, S. and W. T. Tucker (2006), "Sensitivity Analysis Using Probability Bounding," *Reliability Engineering and System Safety*, Vol. 91, No. 10-11, pp. 1435-1442.**

- **Ferson, S., W. L. Oberkampf, and L. Ginzburg (2008), "Model Validation and Predictive Capability for the Thermal Challenge Problem," *Computer Methods in Applied Mechanics and Engineering*, Vol. 197, pp. 2408-2430.**

- **Green, L. L. (2007), "Uncertainty Analysis of Historical Hurricane Data," American Institute of Aeronautics and Astronautics, Paper 2007-1101.**
- **Helton, J. C., J. D. Johnson, C. J. Sallaberry, and C. B. Storlie (2006), "Survey of Sampling-Based Methods for Uncertainty and Sensitivity Analysis," Reliability Engineering and System Safety, Vol. 91, No. 10-11, pp. 1175-1209.**
- **Hasselman, T. K. (2001), "Quantification of Uncertainty in Structural Dynamic Models," *Journal of Aerospace Engineering*, Vol. 14, No. 4, pp. 158-165.**
- **Hills, R. G. (2006), "Model Validation: Model Parameter and Measurement Uncertainty," *Journal of Heat Transfer*, Vol. 128, No. 4, pp. 339-351.**
- **Hills, R. G. and T. G. Trucano (2002), "Statistical Validation of Engineering and Scientific Models: A Maximum Likelihood Based Metric," Sandia National Laboratories, SAND2001-1783, Albuquerque, NM.**

References (continued)

- Kaplan, S. and B. J. Garrick (1981). "On the Quantitative Definition of Risk." *Risk Analysis. 1(1), 11-27.*
- Kennedy, M. C. and A. O'Hagan (2001), "Bayesian Calibration of Computer Models," *Journal of the Royal Statistical Society Series B - Statistical Methodology*, Vol. 63, No. 3, pp. 425-450.
- Morgan, M. G. and M. Henrion (1990). Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. 1st Ed., Cambridge, UK, Cambridge University Press.
- O'Hagan, A. (2006), "Bayesian Analysis of Computer Code Outputs: A Tutorial," *Reliability Engineering and System Safety*, Vol. 91, No. 10-11, pp. 1290-1300.
- Oberkampf, W. L. and M. F. Barone (2006), "Measures of Agreement Between Computation and Experiment: Validation Metrics," *Journal of Computational Physics*, Vol. 217, No. 1, pp. 5-36; also, Sandia National Laboratories, SAND2005-4302.
- Oberkampf, W. L. and T. G. Trucano (2002), "Verification and Validation in Computational Fluid Dynamics," *Progress in Aerospace Sciences*, Vol. 38, No. 3, pp. 209-272.
- Oberkampf, W. L., T. G. Trucano, and C. Hirsch (2003), "Verification, Validation, and Predictive Capability," Sandia National Laboratories, SAND2003-3769, Albuquerque, NM.
- Oberkampf, W. L. and T. G. Trucano (2002), "Verification and Validation in Computational Fluid Dynamics," *Progress in Aerospace Sciences*, Vol. 38, No. 3, pp. 209-272.

References for Section 7 (continued)

- Oberkampf, W.L. and C. J. Roy (2010), Verification and Validation in Scientific Computing, Cambridge University Press, Cambridge, UK.
- Roy, C. J. and W. L. Oberkampf (2011). "A Comprehensive Framework for Verification, Validation, and Uncertainty Quantification in Scientific Computing." *Computer Methods in Applied Mechanics and Engineering. 200(25-28), 2131-2144.*
- Rutherford, B. M. and K. J. Dowding (2003), "An Approach to Model Validation and Model-Based Prediction--Polyurethane Foam Case Study," Sandia National Laboratories, SAND2003-2336, Albuquerque, NM.
- Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola (2008), Global Sensitivity Analysis: The Primer, Wiley, Hoboken, NJ.
- Sprague, M. A. and T. L. Geers (1999), "Response of Empty and Fluid-Filled, Submerged Spherical Shells to Plane and Spherical, Step-Exponential Acoustic Waves," *Shock and Vibration*, Vol. 6, No. 3, pp. 147-157.
- Stern, F., R. V. Wilson, H. W. Coleman and E. G. Paterson (2001), "Comprehensive Approach to Verification and Validation of CFD Simulations-Part 1: Methodology and Procedures," *Journal of Fluids Engineering*, Vol. 123, No. 4, pp. 793-802.
- Zhang, R. and S. Mahadevan (2003), "Bayesian Methodology for Reliability Model Acceptance," *Reliability Engineering and System Safety*, Vol. 80, No. 1, pp. 95-103.

- Zhang, R. and S. Mahadevan (2003), "Bayesian Methodology for Reliability Model Acceptance," *Reliability Engineering and System Safety*, Vol. 80, No. 1, pp. 95-103.